B3

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : C12Q 1/68 — A1

(11) International Publication Number: **WO 00/47766**

(43) International Publication Date: 17 August 2000 (17.08.00)

(21) International Application Number: PCT/GB00/00356

(22) International Filing Date: 8 February 2000 (08.02.00)

(30) Priority Data:
9902971.2 11 February 1999 (11.02.99) GB

(71) Applicant *(for all designated States except US)*: ASTRAZENECA UK LIMITED [GB/GB]; 15 Stanhope Gate, London WIY 6LN (GB).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: GIBSON, Neil, James [GB/GB]; Gadbrook Park, Northwich, Cheshire CW9 7RA (GB). FOY, Carole, Ann [GB/GB]; Gadbrook Park, Northwich, Cheshire CW9 7RA (GB). HODGSON, Darren, Richard [GB/GB]; Gadbrook Park, Northwich, Cheshire CW9 7RA (GB).

(74) Agent: BILL, Kevin; Global Intellectual Property, AstraZeneca UK Limited, Mereside, Alderley Park, Macclesfield, Cheshire SK10 4TG (GB).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: METHOD FOR DETECTING VARIANT NUCLEOTIDES USING ARMS MULTIPLEX AMPLIFICATION

(57) Abstract

The present invention is directed to a method for detecting the presence or absence of a variant nucleotide at each of a plurality of different target sequence in a nucleic acid containing sample by use of a multiplex amplification method that utilises a plurality of ARMS™-allele specific amplification primers possessing non-amplifiable tails in conjunction with an oligonucleotide array for detection amplification of variant nucleic acid sequences. The invention is of particular use in fingerprinting and genome typing, including single nucleotide polymorphism typing, and clinical diagnosis.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

METHOD FOR DETECTING VARIANT NUCLEOTIDES USING ARMS MULTIPLEX AMPLIFICATION

The present invention is directed to the use of a multiplex amplification method that utilises a plurality of ARMS primers possessing non-amplifiable tails in conjunction with a generic oligonucleotide array for detection amplification of variant nucleic acid sequences.

The invention is of particular use in fingerprinting and genome typing, including single nucleotide polymorphism typing, and clinical diagnosis.

Available methods for the amplification and detection of target nucleic acid sequences include use of the polymerase chain reaction (PCR), for example as described in United States patents nos. 4683195 and 4683202.

A significant improvement on the above amplification and detection methods is the Amplification Refractory Mutation System (ARMS™) as claimed in European Patent no. 0 332 435 and corresponding US Patent No. 5,595,890. ARMS is a simple, reliable and rapid PCR-based method for the detection of any single base changes or small deletions, such as mutations or polymorphisms. ARMS detection is particularly useful if the target nucleic acid is an allele in low copy number compared to the alternate allele, for example with somatic underrepresented sequences a single mutant sequence may only be present amongst a background of 10,000 or more wild type sequences.

European Patent EP-B-0416817 discloses non-amplifiable tailed PCR primers which may be adapted for use in the method of the present invention. This disclosure does not exemplify multiplex PCR, using normal or ARMS primers, nor does it teach the added and surprising benefit of enhanced ARMS primer binding specificity observed by the present inventors with large multiplex formats. It also does not demonstrate use of capture oligonucleotides that have minimal secondary structure and/or substantially the same T_m for facilitating and optimising product capture.

Multiplex analysis, whereby numerous distinct target sequences are amplified simultaneously in one PCR reaction by inclusion of a pair of amplification primers for each target sequence, has been reported. Gibbs et al. (Genomics. 7:235-244, 1990) demonstrate multiplex PCR amplification of eight fragments of the hypoxanthine phosphoribosyltransferase (HPRT) gene simultaneously. European Patent Application No. 0 364 255 discloses multiplex PCR amplification of genomic DNA for deletion detection. Results from these PCR and ARMS amplification tests are usually assessed by agarose gel

electrophoresis. The presence or absence, or size change of the expected amplification product(s) determining the status of the target sequence(s). Although it may be desirable to simultaneously analyse samples for as many target sequences as possible, for example, when SNP mapping of individuals, or testing individuals for different disease-linked gene mutations, there are practical limitations to the number of bands (reaction products) that can be analysed on an agarose gel (generally less than about 10). This therefore imposes practical limits on the number of targets that can be assessed in a multiplex amplification reaction. The ability to discriminate many reaction products simultaneously would be of general benefit.

The present inventors have successfully demonstrated a method for simultaneously analysing multiple target sequences, that utilises ARMS primers with non-amplifiable tails, a multiplex amplification reaction and a generic oligonucleotide microarray for capture discrimination of each reaction product. Discrimination arises from capture of each amplification product onto the microarray at a pre-determined and distinct location via its unique tail sequence. Surprisingly, the inventors have found that target amplification selectivity is enhanced in larger multiplex systems, i.e. when more amplification primers are present. The diminution or elimination of false positives, due for example to mis-priming reactions, is extremely important in any amplification detection system.

According to a first aspect of the invention there is provided a method for simultaneously detecting the presence or absence of a variant nucleotide at, or within, each of a plurality of different target sequence in a sample, which method comprises:-

(i) treating the sample, together or sequentially with appropriate nucleoside triphosphates, an agent for polymerisation of the nucleoside triphosphates and a plurality of primers, under suitable hybridisation and then primer extension conditions; each primer having a 3'-end portion substantially complementary to a distinct target nucleic acid sequence which may be present in the sample and a 5'-end portion complementary to one of a set of oligonucleotides immobilised onto a solid surface and optionally, an amplification blocking moiety interposed between said 3'-end and 5'-end portions, the terminal nucleotide of the 3' end portion being either complementary to a suspected variant nucleotide or to the corresponding normal nucleotide, whereby an extension product of the primer is synthesised when the said terminal nucleotide of the primer is complementary to the corresponding nucleotide in the target sequence, no extension product being synthesised when the said terminal nucleotide of the primer is not complementary to the

corresponding nucleotide in the target sequence;

(ii) bringing the extension products produced into contact with a solid surface which has immobilised thereon at pre-determined positions distinct oligonucleotides each comprising a sequence complementary to one or other of the 5'-end portions of the primers; and,

(iii) detecting the presence or absence of the suspected variant nucleotide at each target sequence by detecting the presence or absence of an extension product bound at the pre-determined position on the solid support.

The 5'-end portion represents a single stranded tail portion useful for targeting the amplified products onto a solid support. The amplification primers are amplification refractory mutation system (ARMS) primers, as described in EP-B-0332435 and corresponding US Patent No. 5595890, and as described in EP-B-416817, and additionally described in Newton et al. (Nucleic Acids Research. 17(7):2503-2516, 1989).

ARMS is a technique suitable for detecting the presence or absence of variant nucleotides at a particular loci. It is particularly useful therefore, in diagnostic detection of mutated nucleic acid indicative of tumour phenotype, or in the detection of single nucleotide polymorphisms (SNPs). ARMS is a particularly useful technique where the target sequence is present in low copy number or there is a need to discriminate between two or more alleles, as for example in mutation detection. ARMS mutation detection enables the sensitive detection of specific alleles in the presence of an excess of alternate alleles. In this way somatic mutations can be detected in a background of wild type DNA. ARMS can readily be used to detect 1% mutant sequence in a 99% wild-type background and under appropriate conditions may be used to detect down to 0.01% mutant sequence in 99.99% wild-type background, or less. ARMS uses primers that allow amplification in an allele specific manner. Allele specificity is provided by the complementarity of the 3'-terminal base of a primer with its' respective allele. Amplification is inhibited when the 3'-terminal base of the primer is mismatched. This specificity is maintained when Taq DNA polymerase or other suitable enzyme lacking 3' to 5' proof-reading activity (such as Klenow) is used. An ARMS test is specific when the yield of product from the target allele exceeds the threshold of detection of the system in use and the yield of product from the non-target allele is not detectable. As disclosed in EP-B-0332435 the ARMS primers will preferably possess destabilising mismatches incorporated close to the 3'-terminal nucleotide that discriminates between the different alleles, to enhance specific binding and template amplification from the desired

allele target sequence. The nearer to the 3' terminus of the primer that a destabilising mismatch is incorporated, the greater the effect on destabilisation (See also Newton et al. Nucleic Acids Research. 17:2503-2516, 1989).

In a more preferred embodiment, the tailed primer molecules are used in conjunction with another amplification primer that binds downstream of the target region of interest and lies anti-parallel to the tailed-primer so as to amplify the target region of interest according to the polymerase chain reaction. In one embodiment, this second amplification primer also possess an oligonucleotide tail identical to that of the first targeting primer. In another embodiment, the second amplification primer possesses a detectable label, such as a fluorophor or radioisotope to enable the eventual detection of the amplified product on the microarray. Suitable labelling molecules are well known in the art. Alternatively, a suitable label can be incorporated into the amplified product during its synthesis. A suitable amplification reaction can then be performed so as to generate an amplification product.

The term "nucleotide" as used herein can refer to nucleotides present in either DNA or RNA and thus includes nucleotides which incorporate adenine, cytosine, guanine, thymine and uracil as base, the sugar moiety being deoxyribose or ribose. It will be appreciated however that other modified bases capable of base pairing with one of the conventional bases, adenine, cytosine, guanine, thymine and uracil, may be used in the probes or primers employed in the present invention. Such modified bases include for example 8-azaguanine and hypoxanthine.

The term "oligonucleotide" as used herein is defined as a molecule comprised of two or more nucleotides (i.e deoxyribonucleotides or ribonucleotides), preferably more than five. Its exact size will depend on many factors, such as the reaction temperature, salt concentration, the presence of denaturants such as formamide, and the degree of complementarity with the sequence to which the oligonucleotide is intended to hybridise.

The agent for polymerization of the nucleoside triphosphates may be any compound or system which will function to accomplish the synthesis of primer extension products, including enzymes. Suitable enzymes for this purpose include, for example, *E.coli* DNA Polymerase I, Klenow fragment of *E.coli* DNA polymerase I, T4 DNA polymerase, other available DNA polymerases, reverse transcriptase, and other enzymes, including thermostable enzymes. Because the ARMS™ system relies on binding or non-binding of a 3' terminal nucleotide of the amplification primer, to prevent unmatched 3' primer nucleotides from being

cleaved with subsequent primer extension from the adjacent complementary nucleotide, the agent for polymerisation used must lack 3' to 5' proof-reading activity. The term "thermostable enzyme" as used herein refers to an enzyme which is resistant to heat and catalyzes (facilitates) combination of the nucleotides in the proper manner to form the primer extension. Any convenient template dependent polymerase may be used, however, a thermostable polymerase enzyme such as Taq™, Taq Gold™, or other suitable enzyme lacking 3' to 5' proof-reading activity is particularly preferred.

Similarly any convenient nucleoside triphosphates for conventional base pairing may be used. If required these may be modified for fluorescence. As these may affect polymerisation rates however, for best results, the fluorescently labelled dNTPs are admixed with an excess of wild-type dNTPs, for example in an admixture of between about 1:3 and 1:20.

Further details of convenient polymerases, nucleoside triphosphates, other PCR reagents, primer design, instruments and consumables are given in "PCR" by C.R. Newton and A. Graham (The Introduction to Biotechniques series, Second Edition 1997, ISBN 1 85996 011 1, Bios Scientific Publishers Limited, Oxford). Further guidance may be found in "Laboratory protocols for mutation detection" edited by Ulf Landegren, published by the Oxford University Press, Oxford, 1996, ISBN 0 19 857795 8.on products which are complementary to each nucleic acid strand.

Generally, all four nucleotide triphosphates are provided in the amplification reaction. However, it will be appreciated that where the process of the present invention is to be used for detecting the presence or absence of a suspected variant nucleotide which is adjacent to a portion of the target amplification sequence which does not contain all four different nucleotides, then an extension product of the primer and, if desired, an extension product of the companion amplification primer may be formed in the presence of only the appropriate corresponding nucleoside triphosphates and all four different nucleoside triphosphates would not be necessary.

The term "complementary to" is used herein in relation to nucleotides to mean a nucleotide which will base pair with another specific nucleotide. Thus adenosine triphosphate is complementary to uridine triphosphate or thymidine triphosphate and guanosine triphosphate is complementary to cytidine triphosphate. It is appreciated that whilst thymidine triphosphate and guanosine triphosphate may base pair under certain circumstances

they are not regarded as complementary for the purposes of this specification. It will also be appreciated that whilst cytosine triphosphate and adenosine triphosphate may base pair under certain circumstances they are not regarded as complementary for the purposes of this specification. The same applies to cytosine triphosphate and uracil triphosphate.

"Precise complementarity" or "perfectly matched" as used herein, is in reference to the duplex that the poly- or oligonucleotide strands make with one another to form a double stranded structure such that every nucleotide in each strand undergoes Watson-Crick base pairing with a nucleotide on the other strand. The term also encompasses the pairing of nucleoside analogues, such as deoxinosine, nucleotides with 2-aminopurine bases, and the like, that may be employed. Conversely, a mismatch in a duplex fails to undergo Watson-Crick bonding.

"Substantially complementary" as used herein, refers to poly- or oligonucleotide molecules (or strands) that, under suitable hybridisation conditions (i.e. with reduced stringency), have sufficient complementarity to specifically anneal together, i.e to the exclusion of all other strands, to form a double stranded structure, but wherein one or other strand has, relative to its partner, a limited number of non-complementary (mismatched) nucleotides that are incapable of undergoing Watson-Crick base pairing with the corresponding nucleotide on the other (partner) strand. In a preferred embodiment the number of mismatch nucleotides does not exceed 20%, more preferably 15%, and still more preferably 10% of the total number of nucleotides in the poly- or oligonucleotide.

The expression "target sequence" as used herein means the particular nucleic acid sequence present in the test sample to which the primer is intended to bind, which sequence contains a potential variant nucleotide, the presence or absence of which is to be detected. Thus for example in a single test for p53 mutations that may for example be linked to tumour formation, the p53 gene may contain up to or more than 100, for example 50 target sequences, each target sequence containing a single potential variant nucleotide. It will be appreciated also that there may exist any of three possible variants at any single nucleotide position. It will also be appreciated that the presence of a variant nucleotide at a particular site may be the result of a number of genetic polymorphisms, for example, it may be a simple nucleotide substitution, or, it may be a deletion, an addition, or an inversion, and the like. It will be appreciated therefore that whilst the method of the present invention is of particular interest in detecting the presence or absence of point mutations, the method is equally

applicable to detecting the presence or absence of deletions, including deletions of more than one nucleotide as well as to detecting the presence or absence of substitutions of more than one nucleotide or inversions of more than one nucleotide. In this regard it is simply necessary to know the relevant nucleotides, especially the relevant terminal nucleotide, so that the necessary diagnostic primer(s) may be designed appropriately.

The expression "target amplification sequence" refers to the target sequence and sequence immediately adjacent thereto. When undertaking a PCR reaction the target amplification sequence refers to the sequence that becomes amplified.

It will be appreciated that any extension product formed may be detected in any convenient form, for example in single or double stranded form.

Amplification of the target sequence can be effected by multiple rounds of primer extension from the primer in order to amplify the target nucleic acid such as up to 5, up to 10, up to 15, up to 20, up to 30, up to 40, up to 50 or more times. However, in a preferred embodiment, each targeting oligonucleotide primer is accompanied by a companion primer which facilitates amplification of the target sequence interposed between the two primers according to amplification procedures such as polymerase chain reaction (PCR) or ligase chain reaction (LCR), well known in the art. This companion primer may itself have a non-amplifiable tail portion or it may contain a suitable label which can be used to detect the presence of the amplification product bound onto the solid support.

When the method of the invention employs an amplification system such as the polymerase chain reaction (PCR), the target binding region of the primer and the 5'-end tail region are advantageously arranged such that the tail region remains single stranded, ie. uncopied. Thus the tail region is non-amplifiable in the amplification products. This facet of primer design is disclosed in our European Patent No. 0 416 817 and corresponding US Patent No. 5525494. In order to prevent polymerase mediated chain extension of the tail portion of the ARMS primer template, a blocking moiety is conveniently interposed between the tail portion (5'-end) and the target template binding sequence portion (3'-end). A preferred blocking moiety is a hexethylene glycol (HEG) monomer. Alternatively the primer tail comprises material such as 2'-O-alkyl RNA which will not permit polymerase mediated replication by enzymes lacking reverse transcriptase activity and thus will block the enzyme preventing it from creating a double-strand of the tail region. An alternative way of preventing the tail region (the 5' portion of the primer) from acting as template for the

polymerase enzyme is to link the 3' portion (target binding region) and the 5' portion (tail region) of the primer in the opposite sense to one another, the 3' end being generally in the 5' → 3' sense and the 5' end of the polynucleotide being in the 3' → 5' sense, with linkage via their 5' termini. The presence of the nucleotide sequence of the 5' portion in the opposite orientation prevents the polymerase enzyme from making a fully double stranded amplification product. The 5' portion of the polynucleotide thus becomes a single stranded tail on the amplification product. This single stranded tail can then be utilised for capture of the amplification product onto a complementary capture oligonucleotide attached to a solid support. The more preferred means of blocking the polymerisation agent however, is to incorporate a blocking moiety between the 5' portion and 3' portion of the targeting polynucleotide.

The term " blocking moiety" as used herein means any moiety which when linked, for example covalently linked, between the 3' portion and 5' portion of the polynucleotide is effective to inhibit and preferably prevent, more preferably completely prevent amplification (which term includes any detectable copying) beyond the polymerisation blocking moiety, thus leaving the amplification product with a single stranded tail which is the 5' portion of the polynucleotide. A wide range of blocking moieties may be envisaged for this purpose. For example the polymerisation blocking moiety may comprise a bead, for example a polystyrene, glass or polyacrylamide bead or the polymerisation blocking moiety may comprise a transition metal such as for example iron, chromium, cobalt or nickel (for example in the form of a transition metal complex with the oligonucleotide tail and the target binding nucleotide moiety) or an element capable of substituting phosphorus such as for example arsenic, antimony or bismuth linked between the oligonucleotide tail and the target binding nucleotide moiety. The blocking moiety might similarly involve substitution of the usual phosphate linking groups, for example where oxygen is replaced, leading to *inter alia* phosphorodithioates, phosphorothioates, methylphosphonates, phosphoramidates such as phosphormorpholidates, or other residues known *per se*. Alternative blocking moieties include any 3'-deoxynucleotide not recognised by restriction endonucleases and seco nucleotides which have no 2'-3' bond in the sugar ring and are also not recognised by restriction endonucleases. Newton et al. (Nucleic Acids Research. 21(5):1155-1162, 1993) and EP-B-416817, describe tailed primers with blocking moieties, interposed between the tail

and the target binding portion of the primer, that can be incorporated into the polynucleotides of this invention.

As the tailed ARMS primers may be individually synthesised on a standard oligonucleotide synthesiser, it is preferred that the blocking moiety should be easily incorporated by solid phase oligonucleotide chemistry and should also form a substrate for further extension in the same chemistry. Accordingly, convenient examples of suitable blocking moieties include hexethylene glycol (HEG) and tetraethylene glycol (TEG) phosphoramidites.

The primer is preferably single stranded for maximum efficiency in amplification, but may alternatively be double stranded. If double stranded, the primer is first treated to separate its strands before being used to prepare extension products. Preferably, the primer is an oligodeoxyribonucleotide. The 3' portion of the tailed ARMS primer represents the target binding sequence. This sequence must be sufficiently long to prime the synthesis of extension products in the presence of the agent for polymerization. This sequence can be of any length, depending on the complexity of the target sequence, although it will typically contain between 8 and 60 nucleotides in length such as 10-40 nucleotides, 15-30 nucleotides, particularly 20-30, 17-22, 16-23 or 15-24 nucleotides capable of hybridisation to the target nucleotide sequence, although they may contain more or fewer such nucleotides. Each of the above ranges is a separate and independent embodiment of the invention. Primers having only short sequences capable of hybridisation to the target nucleotide sequence generally require lower temperatures to form sufficiently stable hybrid complexes with the template. The 3' end target binding portion of the tailed ARMS primer sequence need not possess precise complementarity to the target sequence however, it must have sufficient complementarity (be substantially complementary) to bind specifically to the target sequence, that is to say under appropriate hybridisation stringency conditions the target binding region of the primer will hybridise to the target region (if present in the sample) to the exclusion of other regions. It may be advantageous to incorporate nucleotide mismatches into the target binding portion of the ARMS primer in order to assist in destabilising primer binding to incorrect target sequences. The presence of certain mismatches need not however, prevent primer binding to the desired target template sequence.

It is known in the art that in certain circumstances synthesis of a primer extension product might be induced to occur even in the presence of a non-complementary 3'- terminal

residue. This artefactual result may arise from the use of too low a temperature in which case the temperature may be increased, too long a time of incubation/annealing in which case the time may be reduced, too high a salt concentration in which case the salt concentration may be reduced, too high an enzyme concentration, too high a nucleoside triphosphate concentration, an incorrect pH or an incorrect length of oligonucleotide primer. All of these factors are discussed in European Patent Publication No 237,362. A major source of artefactual products is probably allowing the reaction temperature to fall too low, thus permitting too low a stringency, for example by removing the reaction mixture from the heat cycling means, even briefly for example to add the agent for polymerisation (eg. Taq polymerase) especially in the first reaction cycle. In addition to the above artefactual results may also arise from use of a primer which is particularly rich in G (guanine) and C (cytosine) residues. A primer may give rise to difficulty in this regard if it is G/C rich as a whole or particularly if it is G/C rich at its relevant, normally, 3' end. Moreover the precise nature of the base pairing in the region of the 3' end of the primer when in use may be the cause of an artefactual result. Thus the presence of As (adenine) in the base pairing in the region of the 3' end of the primer tends to improve specificity whilst the presence of Gs (guanine) does not. Furthermore the precise nature of the mismatch at the relevant 3' end of the primer may be an important factor in whether or not an artefactual result is obtained. Thus for example an AA or CT mismatch does not normally result in hybridisation, but a GT or AC mismatch may result in a sufficient degree of hybridisation to result in the formation of artefactual product(s). Artefactual results may be avoided by deliberately introducing one or more further mismatched residues, or if desired, deletions or insertions, within the diagnostic primer to destabilise the primer by further reducing the binding during hybridisation. Thus for example any one or more of the 10, for example 6 nucleotides adjacent to the terminal mismatch may be altered to introduce further mismatching. In general only one mismatch in addition to the terminal mismatch may be necessary, positioned for example, 1, 2 or 3 bases from the terminal mismatch. Thus, for example, in relation to the determination of the presence of a normal homozygote, heterozygote or affected homozygote in respect of the Z allele of the α1 anti-trypsin gene we have found that good results may be obtained if the third nucleotide from the 3' terminal nucleotide is altered to generate a mismatch in use. Thus for example we have found that the presence of a C instead of an A as the third nucleotide from the 3' terminus of the diagnostic primer enables normal homozygotes, heterozygotes and affected homozygotes in respect of

the Z allele to be readily distinguished. The best design of diagnostic primer may thus be determined by straightforward experimentation based on the above criteria, such experimentation being well within the ability of the skilled molecular biologist. One feature of the present method is the surprising finding that the incidence of artefactual results, generating false positives, is reduced when a multiplex ARMS assay is performed. Presumably the presence of other competing primers enhances the specificity that the correct primer has for its template.

To avoid false positive or false negative detection it is also essential that none of the tails (the 5' portions of the primers) nor spacer moieties are capable of binding to any nucleic acid in the nucleic acid sample. If the capture hybridisation is to be effected in the presence of all the test sample nucleic acid, i.e. without separation of amplified products from unreacted primers and sample nucleic acid, it is also desirable that none of the capture oligonucleotides on the solid support are capable of binding to any target nucleic acid in the original test sample.

The target binding region of the primer hybridises to template nucleic acid from a sample. The region is of any convenient design available to the person of ordinary skill and is limited only by practical considerations. It may be DNA, RNA or other provided that it provides a substrate for polymerase mediated primer extension. Template binding can be effected at any desired stringency, that is to say under appropriate hybridisation stringency conditions the template binding region of the primer may hybridise to the template region (if present in the template) to the exclusion of other regions. Alternatively template binding may be effected at reduced stringency to extend the primer on any convenient number of related template sequences, such as for example human leukocyte antigen (HLA) genes, or other conserved genes, particularly bacterial or ribosomal RNA genes.

Part of the invention lies in the ability of the primers to specifically distinguish sequences that differ only in the variant nucleotide. The inventors have surprisingly found that as the degree of multiplexing increases (i.e the number of primer pairs) so does the specificity that each primer has for its target sequence, thus reducing the incidence of false-priming. Thus, in distinct embodiments, the method of the invention utilises at least 20, at least 40, at least 50, at least 60, at least 80, at least 100, at least 200, at least 500, or more distinct primers. The preferred range is between about 50 and 200 primers. A suitable format for diagnostic purposes is to utilise suitably adapted microtitre plates (such as 96-well plate)

each well possessing a targeting oligonucleotide capable of capturing one of the ARMS primer tails.

In a preferred embodiment, prior to capturing the amplified products onto the solid surface, the amplification reaction solution is treated to remove those primers that have not participated in primer extension or target template amplification. This may conveniently be performed using techniques well known in the art such as column chromatography using a column substrate suitable to discriminate between oligonucleotides of the particular primer length adopted from predominantly double-stranded target amplified products, or alternatively, by ethanol precipitation or gel filtration.

The solid surface with oligonucleotides immobilised thereon at pre-determined positions is more commonly known in the art as an oligonucleotide microarray.

Microarray (also termed hybridisation array, gene array or gene chip) technology wherein nucleic acid molecules attached to solid substrates at predefined locations in small areas and at high density are used, in conjunction with hybridisation reactions, for identifying and discriminating target nucleic acid sequences, has advanced rapidly in the past few years. These chips or microarrays allow massive parallel data acquisition and are used, for example, in polymorphism detection, clinical mutation detection, expression monitoring, fingerprinting and sequencing. For example, oligonucleotide DNA arrays consisting of short oligonucleotides (e.g. typically 8mers to 20 mers) bound via their 3' termini to a solid surface such as glass or a silicon wafer have been proposed as tools for mutation detection or for the resequencing of genes (e.g. Chee et. al., Science, 1996, Vol. 274, pp 610 - 614; Drobyshev et. al., Gene, 1997, vol.188, pp 45-52).

Methods for preparing a plurality of oligonucleotide sequences and for attaching these to solid supports at high density are known in the art. For example, US Patent No. 4,562,157 describes a method of using photo-activateable cross-linking groups to immobilise pre-synthesised ligands on surfaces. Fodor et al. (Nature. 364:555-556, 1993) and US Patent No. 5,143,854 describe the 'light-directed chemical synthesis' method for synthesising ligands, including oligonucleotides, directly onto a substrate surface at the desired location. US 5,700,637 also describes methods for *in situ* synthesis of oligonucleotides on solid support surfaces. In addition, such methods for preparing microarrays can easily be automated. International Publication No. WO 95/35505 discloses an automated capillary dispensing device and method for applying biological macromolecules to solid supports. International

Publication No. WO 97/44134 also describes devices for delivery of small volumes of liquid (which may contain biological macromolecules) in a precise manner to produce microsized spots on a solid surface to generate a microarray. Similarly, International Publication No. WO 98/10858 also describes an apparatus for the automated synthesis of molecular arrays.

Techniques exist for applying the oligonucleotides to the array at high density and for example, techniques exist for applying well in excess of 10^3 distinct oligonucleotides per 1 cm^2. Many of the advances in microarray technology concern increasing miniaturisation. Aside from the ease in handling and manipulating smaller hybridisation matrices, one significant advantage that smaller chips, with higher density of capture probe, have over larger formats is that the sample does not have to be "stretched out". The technology is unlikely to be widely accepted in the clinical diagnostic market however until costs have been substantially reduced from their current levels.

The design and construction of solid support is well known in the art. Essentially, any conceivable solid substrate may be employed in the invention. A suitable substrate is a material having a rigid or semi-rigid surface, generally insoluble in a solvent of interest such as water. Specific suitable substrates are glass, plastics, polymers, polysaccharides, resins, metal, silica or silica-based material, nylon or nitrocellulose filters, and the like. The solid support may comprise a single sheet of a suitable material such as glass, silicon or plastic so that the pre-selected oligonucleotides are positioned at pre-defined sites based on each oligonucleotide having a distinct and distinguishable set of x,y co-ordinates. Alternatively the solid support may comprise a set of beads of a suitable material such as glass or plastic so that the pre-selected oligonucleotides are positioned at pre-defined sites based on each oligonucleotide residing on a distinct bead. In a preferred embodiment the substrate and/or its surface will be flat glass or single-crystal silicon. Suitable examples of existing laboratory materials that can be utilised are glass microscope slides and microtitre (such as 96-well) plates. The surface of the substrate will preferably contain reactive groups such as carboxyl, amino, hydroxyl, or the like. A polycationic polymer such as polylysine is particularly useful. Most preferably, with fluorescence detection, the surface is non-fluorescent at the wavelength that the analysis is to be performed. The surface of the substrate is also preferably provided with a layer of cross-linking groups to assist attachment of the oligonucleotides to the support. These cross-linking groups will preferably be of sufficient length to permit the oligonucleotides attached to interact freely with their binding partners in solution.

Crosslinking groups may be selected from any suitable class of compounds, for example, aryl acetylenes, ethylene glycol oligomers containing 2-1 monomer units, diamines, diacids, amino acids, and the like. The cross-linking groups may be attached by a variety of methods which are readily apparent to the person skilled in the art. For example, by esterification or amidation reactions of an activated ester of the linking group with a reactive hydroxyl or amine group on the free end of the cross-linking group.

In any diagnostic method relying on hybridisation, precise hybridisation conditions must be adopted on each separate occasion in order to ensure reproducibility of test results. Deviation from optimum recommended hybridisation conditions may have detrimental effects on hybridisation rates of the different amplified products potentially leading to misinterpretation of results. Although not essential, in a preferred embodiment, each of the 5'-end tail sequences of the ARMS primers duplexed to their capture oligonucleotides on the microarray possess substantially the same T_m (for example, within 15°C of the median value, preferably within 8°C of the median value, more preferably within 4°C of the median value), such that under any hybridisation conditions adopted, all of the oligonucleotides on the solid support (herein referred to as "capture oligonucleotides") exhibit approximately the same hybridisation stability with their cognate complementary sequence as the other pairs of oligonucleotide and complementary target sequence. This ensures that approximately equivalent amounts of target DNA are bound to the complementary oligonucleotide on the array at any particular time, facilitating quantitative analysis. As the T_m of all the duplexed oligonucleotides (tails + capture oligos) will be substantially the same however, the optimum temperature for hybridisation can be adopted. In an alternative embodiment, the capture oligonucleotides immobilised on the microarray are designed so as to have minimal secondary structure (such as hairpin loop structure) that might impede hybridisation. With oligonucleotide hybridisation, the optimum hybridisation temperature is generally carried out under conditions that are 5-10°C below the T_m, with the hybridisation and subsequent washes carried out under stringent conditions. Ideally, the hybridisation temperature is controlled precisely, preferably to ±2°C , more preferably to ±0.5°C or better, particularly when the hybridisable length of the capture oligonucleotides are small and there is a need to discriminate between two sequences that may only differ by a single nucleotide at one or other of the termini of the hybridisable sequence. In the case where all of the capture oligonucleotides have substantially the same T_m, it will be apparent that the capture

oligonucleotides need not all be of the same length. Depending on their relative G-C content, two oligonucleotides of different lengths may nevertheless, have the same T_m. Capture oligonucleotides of substantially the same length and G-C content are preferred however.

With regard to the meaning of "substantially the same melting temperature (T_m).", each single-stranded oligonucleotide immobilised on the solid support has, in increasing order of preference, a T_m when bound to its complementary sequence, within 8°C, 7°C, 6°C, 5°C, 4°C, 3°C, 2°C, 1°C, and 0.5°C of the average T_m of all the oligonucleotides immobilised on the solid support. In a more preferred embodiment the oligonucleotides immobilised on the solid support will possess melting temperatures within a range of 0 to 8°C of each other. In a more preferred embodiment at least 90% of the oligonucleotides immobilised on the solid support will possess melting temperatures within 4°C, more preferably 2°C, of each other. In an even more preferred embodiment the oligonucleotides immobilised on the solid support will possess melting temperatures within a range of 0 to 2°C of each other. In the most preferred embodiment, 90-100% of all the oligonucleotides immobilised on the solid support will possess the same melting temperature, and the remaining oligonucleotides will preferably possess melting temperatures within a range of 0 to 2°C of this mode value.

Although it is preferred that all of the oligonucleotides on the solid support fall within the ranges or values for melting temperature as defined above, it is envisaged that a small number of oligonucleotides preferably less than 1-5%, more preferably less than 2% of the total number of oligonucleotides may fall outside these ranges or values.

With regard to the meaning of "substantially the same length", in increasing order of preference, the length of each capture portion of the oligonucleotide immobilised on the solid support will be within or equal to16, 12, 10, 8, 6, 5, 4, 3, 2, and 1 nucleotide(s) of the average length of all the capture portions of the oligonucleotides immobilised on the solid support. In a preferred embodiment the oligonucleotide capture portions will each be of a length that is within 0-8 nucleotides of each other.

With regard to the meaning of "substantially the same G-C content", in increasing order of preference, each of the oligonucleotides immobilised on the solid support will have a G-C content within or equal to 25%, 20%, 15%, 10%, 10%, 5% and 2% of the average G-C content of all the immobilised oligonucleotides. In a preferred embodiment, 95-100% of all the oligonucleotides immobilised on the support will have a percentage G-C content within 10% of the median value and the remainder will preferably be within 25% of the median

value. More preferably the percentage G-C content of the oligonucleotides immobilised on the solid support will be within 8% of each other.

In a preferred embodiment the capture portion of all of the oligonucleotides are of the same length and have the same G-C content.

Although it is preferred that all of the oligonucleotides on the solid support fall within the ranges for length and G-C content as defined above, it is envisaged that a small number of oligonucleotides preferably less than 1-5%, more preferably less than 2% of the total number of oligonucleotides may fall outside these ranges. The presence of oligonucleotides on the solid support that possess capture sequences that fall outside the preferred sequence composition (i.e. length and G-C content) need not diminish the utility of the generic array particularly if these capture sequences are excluded from being used as capture molecules when the array is in use.

The melting temperature (T_m) referred to herein, is defined as the temperature at which duplex DNA exists in a ratio of 50:50 in hybridised and dissociated form under equilibrium conditions. The principal governing factors determining T_m are sequence length and G-C content. The theoretical and experimental procedure for determining the T_m is disclosed in Molecular Cloning-A Laboratory Manual, Second Edition, J Sambrook et al., Cold Spring Harbor, Chapter 11 section 46 and 55. In essence, for oligonucleotides shorter than 18 nucleotides, the T_m of the hybrid is estimated by multiplying the number of A + T residues in the hybrid by 2°C and the number of G + C residues by 4°C and adding the two together. For oligonucleotides between approximately 14 and 70 nucleotides in length, the following equation devised by Bolton and McCarthy, (P.N.A.S. 48:1390, 1962) for determining T_m of long DNA molecules is also applicable:

$$T_m = 81.5 - 16.6(\log_{10}[Na^+]) + 0.41(\% G + C) - (600/N).$$

Wherein N = chain length and $[Na^+]$ is the ionic strength of the hybridisation solution.

The capture oligonucleotides attached to the solid support, and thus also the 5'end tail regions of the ARMS primers, generally have a hybridisable sequence between 5 and 100 nucleotides in length. The preferred length of hybridisable sequence is in the range of 10 - 50, more preferred is 12-20. The hybridisable sequence is that portion of the capture oligonucleotide (capture portion) that is designed and available for hybrid formation with its complementary sequence. As used herein in reference to the hybridisable sequence and capture portion are used interchangeably. Non-hybridisable sequence of the capture

oligonucleotide might represent flanking or tether sequences. Tether sequences not only serve to anchor the oligonucleotide to the solid support but also serve to distance the hybridisable portion of the capture oligonucleotide from the solid support to alleviate steric interference.

The capture oligonucleotide molecules may be individually synthesised on a standard oligonucleotide synthesiser. These oligonucleotide (oligos) may then be attached to the substrate matrix by any of a variety of techniques known in the art such as by using photochemical reagents, such as disclosed in US Patent No. 4,542,102 and 4,713,326. US Patent No. 4,562,157 also describes a method of using photo-activatable cross-linking groups to immobilise pre-synthesised ligands on surfaces. Alternatively, the oligonucleotides can be synthesised directly onto the solid surface using photolithography techniques, such as disclosed in US Patent No. 5,143,854, or other methods such as disclosed in US Patent No. 5,700,637, or International Publication No's: WO 95/35505, WO 97/44134 or WO 98/10858. Schena et al. (TIBTECH 16(7):301-306, 1998) reviews the recent advances in microarray technology including the various means of constructing these arrays.

In order to avoid or alleviate steric factors during the capture hybridisation reaction, it may be desirable to use a tether/linker molecule to tether the capture oligonucleotides to the solid support. Shchepinov et al. (N.A.R. 25:1155-1161, 1997) disclose the use of various amino group-containing phosphoramidite moieties to distance the capture oligonucleotide from the solid support and thus alleviate steric interference. They found that with a linker of at least 40 atoms in length they obtained up to 150-fold increased hybridisation yields. Based on the teaching in Shchepinov et al., the person skilled in the art would be able to design and synthesise suitable tether/linker molecules to reduce steric interference of the support on hybridisation behaviour of the immobilised capture oligonucleotides of the invention.

An advantage of using non-amplifiable single-stranded tailed primers in the method of the invention, is that the target amplified product has a single-stranded portion which can be hybridised without denaturation to the solid support containing the immobilised pre-selected capture oligonucleotide sequences. Current array technology requires the target nucleic acids to be denatured or rendered single-stranded some other way, prior to capture on the array.

The oligonucleotide microarray suitable for capturing the tailed ARMS primers used in the multiplex amplification method of the present invention is another aspect of the invention.

The multiplex tailed ARMS primers amplification with oligonucleotide microarray capture method of the invention is useful in any setting where it is desirable to identify the presence of one or more specific nucleic acid sequences from a population of sequences. Examples of uses are fingerprinting, diagnostic identification, genotyping of organisms (including single nucleotide polymorphism detection) and environmental monitoring. A particular use is in the detection of inherited disease or acquired disease, such as tumours.

Any population of nucleic acids represents a suitable test sample. Sources of test sample nucleic acid include human cells such as circulating blood, buccal epithelial cells, cultured cells and tumour cells. Other mammalian tissues and cultured cells are also suitable sources of template nucleic acids. In addition, viruses, bacteriophage, bacteria, fungi and other micro-organisms can be the source of nucleic acid for analysis. The DNA may be genomic or it may be cloned in plasmids, bacteriophage, bacterial artificial chromosomes (BACs), yeast artificial chromosomes (YACs) or other vectors. RNA may be isolated directly from the relevant cells or it may be produced by *in vitro* priming from a suitable RNA promoter or by *in vitro* transcription.

In general, the detection of allelic variation requires a mutation discrimination technique, optionally an amplification reaction and a signal generation system. Table 1 lists a number of conventional signal generation systems that may be employed to detect the tailed ARMS products generated according to the invention.

Table 1 - Signal Generation or Detection Systems

Fluorescence: Fluorescence intensity, Fluorescence resonance energy transfer (FRET), Fluorescence quenching, Fluorescence polarisation - United Kingdom Patent No. 2228998.

Other: Chemiluminescence, Electrochemiluminescence, Raman, Radioactivity, Colorimetric, Hybridisation protection assay, Mass spectrometry.

One way of detectably labelling the amplified products is to incorporate a suitable label into the tailed ARMS primers and/or the companion amplification primer, or to incorporate a label such as fluorescein-dUTP into the amplified products. In essence, any number of conventional detectable markers such as radioisotopes, fluorescent labels, chemiluminescent compounds, labelled binding proteins, magnetic labels, spectroscopic markers and linked enzymes might be used. The labelling technique employed is a matter of

personal choice for the person skilled in the art. Fluorescent labels are preferred because they are less hazardous than radiolabels, they provide a strong signal with low background and various different fluorophors capable of absorbing light at different wavelengths and/or giving off different colour signals exist to enable comparative analysis in the same analysis. For example, fluorescein gives off a green colour, rhodamine gives off a red colour and both together give off a yellow colour. If the tailed ARMS primers are labelled, the unextended ARMS primers can be separated from extended primers prior to the capture step using conventional techniques such as column chromatography, ethanol precipitation or gel filtration. Other potential means for discriminating between those captured primers that have been primer extended and thus form part of an amplified product and those that have not been primer extended include the use of intercalating agents (i.e dyes such as ethidium bromide) that become incorporated into duplex nucleic acid or the use of labelled binding proteins or antibodies or other reagents that recognise helix formation (such as the target nucleic acid/targeting oligonucleotide hybrid), see for example US Patent No. 4,582,789, or the use of a ligand binding to the minor groove such as Hoechst 33258 fluorescent dye or the use of ligands which recognise the minor groove of DNA in a sequence specific manner, see for example, "Recognition of the Four Watson-Crick Base Pairs in the DNA Minor Groove by Synthetic Ligands." S. White, J. W. Szewczyk, J. M. Turner, E. E. Baird and P. B. Dervan, Nature, 391, 468 (1998). Convenient intercalators and minor groove binders will be apparent to the person skilled in the art (cf Higuchi et al. BioTechnology. 10:413-417, 1992). In a preferred embodiment of the invention, capture of the hybrid detection product by the oligonucleotide on the solid support is detected using one or more minor groove binding probes.

It will be apparent to the person skilled in the art that there are other conventional detection means that can be employed in order to detect the amplified products captured on the solid support. The essential feature is that hybridisation of the products onto the capture oligonucleotide on the microarray causes a detectable change in a signalling system. Any convenient signalling system may be used. By way of non-limiting example we refer to the measurement of the change in fluorescence polarisation of a fluorescently labelled species (European Patent No. 0 382 433), DNA binding proteins, intercalators, or the incorporation of detectable (modified) dNTPs into the primer extension products or the target nucleic acids.

The detection of specific interactions may be performed by detecting the positions where the labelled target sequences are attached to the array. Radiolabelled probes can be detected using conventional autoradiography techniques. Use of scanning autoradiography with a digitised scanner and suitable software for analysing the results is preferred. Where the label is a fluorescent label, the apparatus described, e.g in International Publication No. WO 90/15070, US Patent No. 5, 143,854 or US Patent No. 5,744,305 may be advantageously applied. Indeed, most array formats use fluorescent readouts to detect labelled capture:target duplex formation. Laser confocal fluorescence microscopy is another technique routinely in use (M.J.Kozal et al., Nature Medicine. 2:753-759, 1996). Mass spectrometry may also be used to detect oligonucleotides bound to a DNA array (Little DP et al, Analytical Chemistry. 69(22):4540-4546, 1997). Whatever the reporter system used, sophisticated gadgetry and software may be required in order to interpret large numbers of readouts into meaningful data (such as described, for example, in US Patent No. 5,800,992 or International Publication No. WO 90/04652).

Further systems include two-component systems where a signal is created or abolished when the two components are brought into close proximity with one another. Alternatively, a signal is created or abolished when the two components are separated following binding of the target binding region.

Both elements of the two component system may be provided on the same or different molecules. By way of example the elements are placed on different molecules, target specific binding displaces one of the molecules into solution leading to a detectable signal. One of the components may be attached to the capture oligonucleotide, or the solid support itself. For example, the array could consist of fluorescein labelled oligonucleotides of, for example 20 residues in length. Prior to addition of the sample, a set of short quencher oligonucleotides (say 10 residues in length) complementary to the array and labelled with DABCYL (or methyl red) could be added to the array. The short complementary DABCYL oligonucleotides bind to the corresponding 'address' on the array and the fluorescence of the fluorescein labels is quenched. The sample is then purified so that unextended primers or unbound probes are separated from extended or bound products. The bound or amplified products are then added to the microarray and the tail portions which are fully complementary to the oligonucleotides on the microarray bind to the microarray with displacement of the quencher oligonucleotides. This results in the microarray oligonucleotides fluorescing as a result of the binding of the

appropriate product. In this format a fluorescent signal is produced by separating two species bound to the array surface. One advantage of this format is that it permits quality control of the array. When the fluorescent array is manufactured it can be scanned in a fluorescent scanner and any defects such as a probe oligonucleotide which has failed to attach to the surface will be detected as a non or weakly fluorescent spot on the array. Efficient quenching when the quencher oligonucleotides are added can also be monitored before the test products are added.

Convenient two-component systems may be based on the use of energy transfer, for example between a fluorophore and a quencher. In a particular aspect of the invention the detection system comprises a fluorophore/quencher pair. Convenient and preferred attachment points for energy transfer partners may be determined by routine experimentation. A number of convenient fluorophore/quencher pairs are detailed in the literature (for example Glazer et al, Current Opinion in Biotechnology. 8:94-102, 1997,) and in catalogues such as those from Molecular Probes, Glen and Applied Biosystems (ABI). Any fluorescent molecule is suitable for signalling provided it may be detected on the instrumentation available. Most preferred are those compatible with the 488 nm line of the Argon ion laser (Fluorescein and Rhodamine derivatives). The quencher must be able to quench the dye in question and this may be via a Fluorescence Resonance Energy Transfer (FRET) mechanism involving a second, receptor fluorophore, or more preferably via a collisional mechanism involving a non-fluorogenic quencher such as DABCYL, which is a "Universal" quencher of fluorescence or methyl red. Furthermore it is preferred that the selected fluorophores and quenchers are incorporated, most conveniently via phosphoramidite chemistry, into the capture oligonucleotides and/or targeting polynucleotides and/or second primer required for example, when undertaking PCR-based amplification. FAM, a fluorescein dye with an excitation optimum at ~490nm, is a convenient donor.

A further embodiment consists of a fluorescently labelled DNA array where the array oligonucleotides are labelled with a fluorophore which either does not fluoresce at the irradiation frequency or is only weakly fluorescent at this frequency. The tailed ARMS primers or products are labelled with a fluorophore which is substantially fluorescent at the irradiation frequency and which forms an energy transfer pair with the fluorophore label on the DNA array. When the purified tailed ARMS products are bound to the array and subjected to irradiation there is energy transfer between the fluorophore on the tailed ARMS

product and that on the array. The array fluorophore at the specific binding sites increase substantially in its fluorescent brightness and this may be detected by scanning the array. This is an example of increasing the fluorescent signal on the array by bringing two species close together by hybridisation.

When capturing the products on the array or when annealing the primers to the template to effect primer extension, hybridisation conditions chosen are designed to be as close as possible to the T_m of the duplexes. The concentration of salt in the hybridisation solution used is particularly significant. At 1M NaCl, G:C base pairs are more stable than A:T base pairs. Similarly, double stranded oligonucleotides with a higher G-C content have a higher T_m than those of the same length but with a higher A-T content. If slight differences, i.e single nucleotide differences, amongst the target nucleic acids need to be distinguished, establishing optimum hybridisation conditions is important, particularly, when the hybridisable length of the oligonucleotides is small (< approximately 30-mers). Where, because of the diverse composition of the target sequences, there is a broad range of T_m, either a less than optimum compromise set of hybridisation conditions could be adopted, or conditions could be manipulated so as to diminish the T_m dependence on nucleotide composition by using chaotropic hybridisation solutions. This can be effected, for example, by incorporation into the hybridisation solution of a tertiary or quaternary amide. Tetramethylammoniumchloride (TMACl) is particularly suitable when used at concentrations of between 2M and 5.5M. A preferred concentration range being 3M - 4M. Compared to the presence of 1M NaCl in the hybridisation solution, use of up to 5M TMACl can enhance hybridisation specificity by up to 40-fold.

Once a particular sequence has, or group of sequences have, been hybridised to the microarray and the pattern of hybridisation analysed, the microarray can be treated to remove the bound sequences in preparation for reuse of the microarray by exposure to a second or subsequent set of target sequences. In order to do this the hybrid duplexes are disrupted and the solid support matrix treated in order to remove all traces of the original target. To effect this, the matrix may be treated with various detergents or solvents to which the substrate, the oligonucleotides and the linkages to the substrate are inert. This treatment may involve an elevated temperature treatment, treatment with organic or inorganic solvents, modifications in pH, and other means for disrupting specific interactions. Examples of methods that could be used are: (1) Washing the array with 50 mM sodium hydroxide to disrupt base pairing by high

pH. (2) Washing the array with pure water and at high temperature (e.g. > 80°C) to disrupt base pairing by high stringency. (3) Addition of oligonucleotide sequences complementary to the tail sequences (and identical to the chip sequences) to disrupt base pairing by exchange with the sequences in free solution. Other methods for disrupting duplex formation are well known in the art (see for example Sambrook et al. *ibid*).

According to a further aspect of the invention there is provided a kit for detecting the presence or absence of one or more variant nucleotides in one or more nucleic acids contained in a sample, which kit comprises:-

(i) a plurality of primers, one primer for each potential variant nucleotide of a target sequence to be detected, each primer having a 3'-end portion substantially complementary to a distinct target nucleic acid sequence which may be present in the sample and a 5'-end portion complementary to one of a set of oligonucleotides immobilised onto a solid surface and optionally, an amplification blocking moiety interposed between said 3'-end and 5'-end portions, the terminal nucleotide of the 3' end portion being either complementary to a suspected variant nucleotide or to the corresponding normal nucleotide, whereby an extension product of the primer is synthesised when the said terminal nucleotide of the primer is complementary to the corresponding nucleotide in the target sequence, no extension product being synthesised when the said terminal nucleotide of the primer is not complementary to the corresponding nucleotide in the target sequence; and

(ii) a solid support having a surface on which is immobilised at pre-determined positions a plurality of distinct oligonucleotides that are each independently complementary to one or other of the 5'-end portions of the primers in (i).

In a preferred embodiment the kit also contains one or more of the four different nucleoside triphosphates and/or an agent for polymerisation of the nucleoside triphosphates and/or instructions for use. In another preferred embodiment the kit comprises a set of at least two primers for each target sequence, the terminal nucleotide of at least one primer being complementary to a suspected variant nucleotide associated with a known genetic disorder or a polymorphism and at least one of the other primers being a companion primer as described hereinbefore. The kit may therefore comprise sets of oligonucleotide primers, each set targeting different alleles at a specific loci. In another preferred embodiment the solid support is a microtitre plate with each capture oligonucleotide immobilised in a distinct well of the plate.

The invention will now be further illustrated by the following non-limiting examples. The examples refer to the following figures, in which:

Figure 1a - Records the absorbance values (OD405nm) of the captured amplification products for Experiment 1 (23-plex). Signals generated from the 10% mutant template are shown as bars and those from 100% wild-type template are shown as diamonds.

Figure 1b - Records the ratio of signals from mutant and wild-type templates (ODm/ODw) from Experiment 1.

Figure 2a - Records the absorbance values (OD405nm) of the captured amplification products for Experiment 2 (84-plex). Signals generated from the 10% mutant template are shown as bars and those from 100% wild-type template are shown as diamonds.

Figure 2b - Records the ratio of signals from mutant and wild-type templates (ODm/ODw) from Experiment 2.

Figure 3a - Records the absorbance values (OD405nm) of the captured amplification products for Experiment 3 (23-plex). Signals generated from the 10% mutant template are shown as bars and those from 100% wild-type template are shown as diamonds.

Figure 3b - Records the ratio of signals from mutant and wild-type templates (ODm/ODw) from Experiment 3.

Figure 4a - Records the absorbance values (OD405nm) of the captured amplification products for Experiment 4 (84-plex). Signals generated from the 10% mutant template are shown as bars and those from 100% wild-type template are shown as diamonds.

Figure 4b - Records the ratio of signals from mutant and wild-type templates (ODm/ODw) from Experiment 4.

Example 1

Multiplex tailed ARMS assay to detect p53 mutations.

Many potential mutation sites in p53 have been identified (P.Hainaut et al, Nucleic Acids Research. 26(1):205-213, 1998).

80 ARMS primers were designed for the specific detection of some of the mutations in exons 5-8 of the p53 tumour suppresser gene (Table 2 lists the 80 codon positions and specific mutations for which ARMS primers were designed and prepared). Uniquely identifying non-amplifiable tails (with T_ms in the range of 53°C to 58°C) with hexaethylene

glycol links between the primer and the tail sequences were added to 19 of these ARMS primers (marked * in Table 2). The ARMS primers were then multiplexed together with 2 reverse primers designed to give PCR products with the ARMS primers specific for mutations in exons 5&6 and 8 respectively. Tailed primer sets which act as control primers for the detection of p53 exons 5&6 and exon 8 sequence were also included.

Table 2.

List of potential p53 mutations on which the ARMS primers were prepared.

132 CAG*	132 AGG	135 TAC	141 TAC	151 TCC	151 ACC*
152 CTG	154 GTC	156 CCC*	157 TTC*	158 CTC*	158 CAC
159 GAC*	159 GTC	161 ACC	163 TGC	173 ATG	173 TTG*
175 CAC*	176 TTC*	177 CGC*	179 CGT	179 TAT	192 TAG
193 CGT	195 ACC	196 TGA	205 TGT	213 TGA	220 TGT*
228 AAC	234 TGC	237 ATA	238 TAT	241 TTC	242 TTC
244 TGC	245 TGC	245 GAC	245 AGC	245 CGC	245 GTC
245 GCC	248 CTG	248 CAG	248 TGG	248 GGG	248 CCG
249AGC	249AGT	249 ACG	249 ATG	249 GGG	249 AAG
258 AAA	258 GGA	266 AGA	266 GAA	266 TGA	266 GTA
272 ATG	273 CCT*	273 CTT*	273 CAT*	273 TGT*	273 AGT
273 GGT	275 TAT	278 CTT	278 TCT	280 ACA*	282 TGG
282 CAG	282 CCG	282 CTG	282 GGG	285 AAG*	286 AAA*
298 TAG*	306 TGA				

* - denotes the mutant codon for which a tailed ARMS primer was prepared.

By way of representative example, the following represents the specific tailed ARMS primer used to detect p53 175 CAC mutation (* denotes the HEG group):

5'-GCTTTATGTCCACAGATTTC*ATACACAGCACATGACGGAGGTTGTGAGCCA-3'

SEQ ID No. 1 represents the 5' tail portion; SEQ ID No. 2 represents the 3' targeting portion

The specific reverse primer used with the above ARMS primer is:

5'-ACCCGGAGGGCCACTGACAAC-3' (SEQ ID No. 3)

Two templates, one containing the 175 CAC p53 mutation and the other containing the 132 CAG mutation in the p53 gene were prepared by primer directed mutagenesis (see Higachi et al. NAR. 16:7350-7367, 1988). This gave templates of 500 - 800 bp containing the mutation of interest. Wild type templates were prepared by PCR amplification of wild-type DNA to give the corresponding 500 - 800 bp fragments.

All templates were quantified relatively by real time analysis on an ABI Prism 7700 using a quantitative PCR reaction. Quantitated cassettes (mutant synthetic templates prepared by site directed mutagenesis) were then used to prepare wild-type/mutant admixtures.

Oligonucleotides of precise complementary sequence to the 19 tail sequences were synthesised with 3' biotin moieties. These capture sequences were bound to the wells of a streptavidin coated microtitre plate (one capture sequence per well).

Reaction conditions: Each ARMS primer was present at 50 nM concentration, the reverse primers at 500mM concentration and wild type dNTPs at 50 μM each. Fluorescein-dUTP was also included at 0.5 μM. The buffer was 50 mM KCl, 10 mM tris, 1.2 mM $MgCl_2$ at pH 8.3. 4 Units of AmpliTaq Gold™ were used per amplification. 10^5 copies of template were added. Cycling conditions were 94°C for 20 minutes then 35 cycles of (94°C for 1 minute, 60°C for 1 minute). In each experiment, three parallel experiments were run using (a) wild type template, (b) mutant/wild-type admixture template and (c) a no-template control.

Following amplification, the PCR products were divided between the capture wells of the microtitre plate. Hybridisation between the PCR products and the capture oligos took place overnight at 55°C in 3M TMAC, 1M Tris (pH 7.5), 0.5M EDTA, 0.01% Triton-X-100, 0.1mg/ml herring sperm DNA . Unbound products were then washed off (2 washes in phosphate buffered saline (PBS)). The PCR products were detected by ELISA detection of incorporated fluorescein-dUTP using an anti-fluorescein-alkaline phosphatase antibody-enzyme conjugate. Colour development was by addition of *p*-nitrophenyl phosphate and the OD 405 was determined after 30 minutes.

Experiment 1 - Detection of a single point mutation at codon position 175 of p53 present as template at a concentration of 10% using a 23-multiplex system.

Admixtures were prepared between template containing p53 mutant codon 175 CAC and wild-type template with the mutant sequence present at 10% of the total.

Three separate amplifications (using the multiplex of 19 tailed ARMS primers, two reverse primers and 2 control template primers) were carried out on: (a) the admixture; (b) wild-type template; and, (c) no-template control. The amplification products were then each added to a separate array consisting of a microtitre dish with 17 capture oligos for a subset of the primers plus 2 capture oligo for the exon 5&6 control reaction immobilised thereon.

Figures 1a shows the ODm and ODw values with the relevant no-template control values subtracted for the mutant (10% 175 CAC) and 100% wild-type templates amplified with a multiplex of 19 mutant primers, 2 reverse primers and 2 control primers for different exons of the p53 gene plus two reverse primers. The signals generated from the mutant template are shown as bars and those from the WT as diamonds. All signals are background corrected. The control primer generates signal from both templates whereas the 175 CAC primer only generates signal from the mutant template. In addition, a further two primers produce signal non-specifically from both templates (132 CAG and 220 TGT).

Figure 1b shows the ratio of signals from the mutant and WT templates (ODm/ODw). The primer specific for the mutation (175 CAC) produces a significant signal from the mutant template only with the ratio being approximately 50. The control signals and inappropriate priming signals are similar on both mutant and WT templates and produce ratios of approximately 1.

Experiment 2.- Detection of a single point mutation at codon position 175 of p53 present as template at a concentration of 10% using a 84-multiplex system.

The following experiment demonstrates that the specificity of allele discrimination increases with increasing numbers of targeting oligonucleotides in the multiplex reaction. Experiment 1 was repeated except that the multiplex was increase to an 84-plex system by the addition of further ARMS primers for the other possible p53 variants listed in Table 2.

The effects of the increase in multiplex size is examined by comparing the new signals (from the 19 tailed primers present in the 84-plex) to the signals from the original 19 tailed

primers in the first example. Again the only primer which produces significantly greater signal on the mutant template is the mutation specific 175 CAC primer (see Figure 2a). Of note is that with the increase in primer number, one of the primers which previously primed erroneously (132 CAG) no longer does so. This increase in specificity is further confirmed by examining the ODm/ODw ratios which for the 175 CAC test has increased to almost 300:1 (Figure 2b).

The greater specificity of primers present in the 84-plex is further seen using the 132 CAG mutation as template (compare experiments 3 and 4 below).

Experiment 3.- Detection of a single point mutation at codon position 132 of p53 present as template at a concentration of 10% using a 23-multiplex system.

Experiment 1 was repeated except that the 175 CAC mutant template was substituted for by the 132 CAG mutant template. In addition, rather than using the 19 capture oligonucleotides as in experiment 1, in this experiment efficient target amplification was demonstrated using just 8 key capture oligonucleotides.

Figures 3a shows the ODm and ODw values with the relevant no-template control values subtracted for the mutant (10% 132 CAG) and 100% wild-type templates amplified with a multiplex of 19 mutant primers, 2 reverse primers and 2 control primers for different exons of the p53 gene. As demonstrated previously the 132 CAG primer primes and extends off WT template in the context of the 23-plex. Although, in this experiment, the signal from the mutant template is greater than that from the WT template, the priming of this primer on the WT template results in an ODm/ODw ratio of only 7 (Figure 3b). A further effect of non-specificity is seen from the 151 ACC test which produces a small signal from the mutant template and generates a ODm/ODw ratio of 25 (Figure 3b).

Experiment 4.- Detection of a single point mutation at codon position 132 of p53 present as template at a concentration of 10% using a 84-multiplex system.

The following experiment demonstrates that the specificity of allele discrimination increases with increasing numbers of targeting oligonucleotides in the multiplex reaction. Experiment 2 was repeated except that the 175 CAC mutant template was substituted for the 132 CAG mutant template. As with experiment 3, only 8 capture oligonucleotides were used to illustrate capture of target amplified products.

This experiment shows the same templates as in Experiment 3, but amplified with the 84 primer multiplex (Figure 4a). Consistent with previous findings the 132 CAG primer no longer primes off WT template resulting in an increase in the ODm/ODw ratio to 250 (Figure 4b). The 151 ACC primer has also stopped priming off the 132 CAG mutant template in the 84-plex.

CLAIMS

1. A method for detecting the presence or absence of a variant nucleotide at each of a plurality of different target sequence in a nucleic acid containing sample, which method comprises:-

(i) treating the sample, together or sequentially with appropriate nucleoside triphosphates, an agent for polymerisation of the nucleoside triphosphates and a plurality of primers, under suitable hybridisation and then primer extension conditions, each primer having a 3'-end portion substantially complementary to a distinct target nucleic acid sequence which may be present in the sample and a 5'-end portion complementary to one of a set of oligonucleotides immobilised onto a solid surface and optionally, an amplification blocking moiety interposed between said 3'-end and 5'-end portions, the terminal nucleotide of the 3' end portion being either complementary to a suspected variant nucleotide or to the corresponding normal nucleotide, whereby an extension product of the primer is synthesised when the said terminal nucleotide of the primer is complementary to the corresponding nucleotide in the target sequence, no extension product being synthesised when the said terminal nucleotide of the primer is not complementary to the corresponding nucleotide in the target sequence;

(ii) bringing the extension products produced into contact with a solid surface which has immobilised thereon at pre-determined positions distinct oligonucleotides each comprising a sequence complementary to one or other of the 5'-end portions of the primers; and,

(iii) detecting the presence or absence of the suspected variant nucleotide at each target sequence by detecting the presence or absence of an extension product bound at the pre-determined position on the solid support.

2. A method as claimed in claim 1 wherein prior to step (ii) the reaction mixture is treated to remove those primers that have not participated in primer extension or target template amplification.

3. A method as claimed in claim 1 or 2 wherein there are at least 20, particularly at least 50, distinct amplification primers.

4. A method as claimed in claim 1 or 2 wherein there are at least 100, particularly at least 200, distinct amplification primers.

5. A method as claimed in any of claims 1-4 wherein each tailed primer molecule is used in conjunction with a companion primer.

6. A method as claimed in any of the preceding claims wherein each of the 5'-end tail sequences of the ARMS primers when duplexed to their capture oligonucleotides on the microarray possess substantially the same T_m.

7. A method as claimed in claim 1 wherein hybridisation of the products generated in step (i) onto the oligonucleotide on the solid surface in step (ii) causes a detectable change in a signalling system.

8. A kit for detecting the presence or absence of one or more variant nucleotides in one or more nucleic acids contained in a sample, which kit comprises:-

(i) a plurality of primers, one primer for each potential variant nucleotide of a target sequence to be detected, each primer having a 3'-end portion substantially complementary to a distinct target nucleic acid sequence which may be present in the sample and a 5'-end portion complementary to one of a set of oligonucleotides immobilised onto a solid surface and optionally, an amplification blocking moiety interposed between said 3'-end and 5'-end portions, the terminal nucleotide of the 3' end portion being either complementary to a suspected variant nucleotide or to the corresponding normal nucleotide, whereby an extension product of the primer is synthesised when the said terminal nucleotide of the primer is complementary to the corresponding nucleotide in the target sequence, no extension product being synthesised when the said terminal nucleotide of the primer is not complementary to the corresponding nucleotide in the target sequence; and

(ii) a solid support having a surface on which is immobilised at pre-determined positions a plurality of distinct oligonucleotides that each independently comprise a sequence complementary to one or other of the 5'-end portions of the primers in (i).

9. A kit as claimed in claim 8 also comprising one or more of:

(i) nucleotide triphosphates;

(ii) a polymerisation agent;

(iii) control DNA; and,

(iv) instructions for use.

Figure 1a.

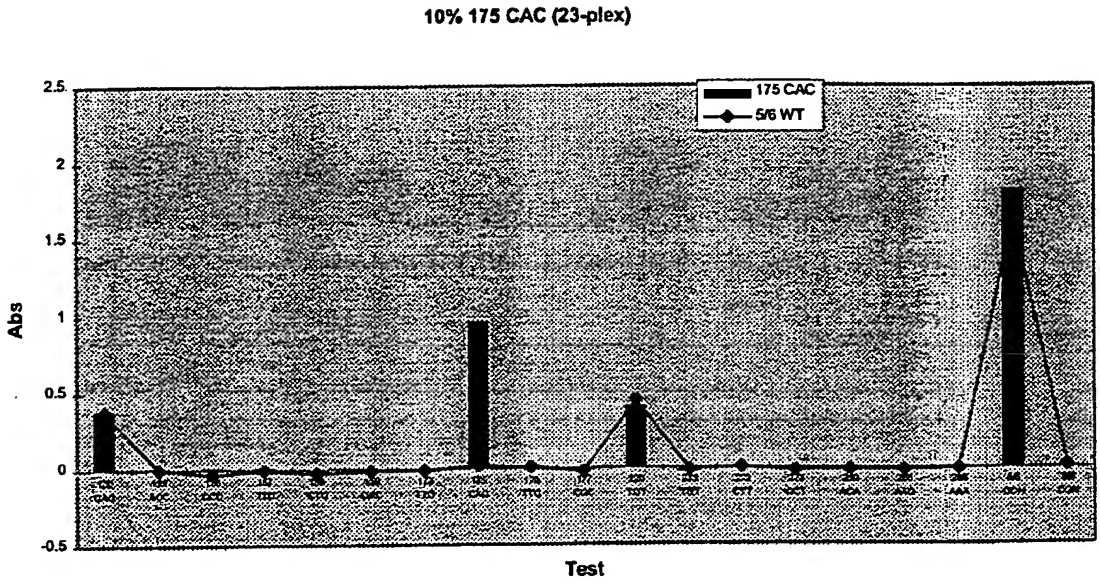


Figure 1b

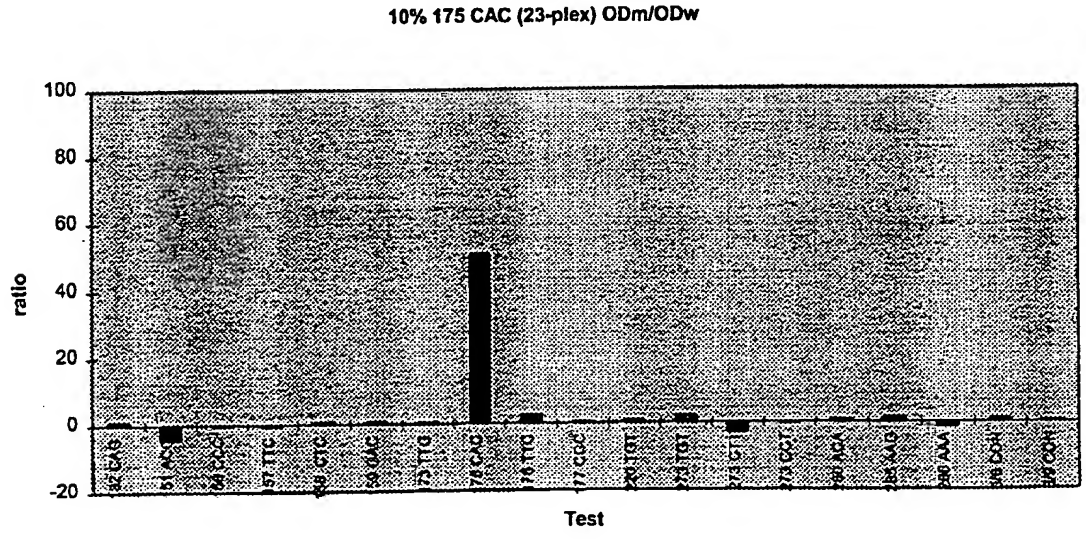


Figure 2a

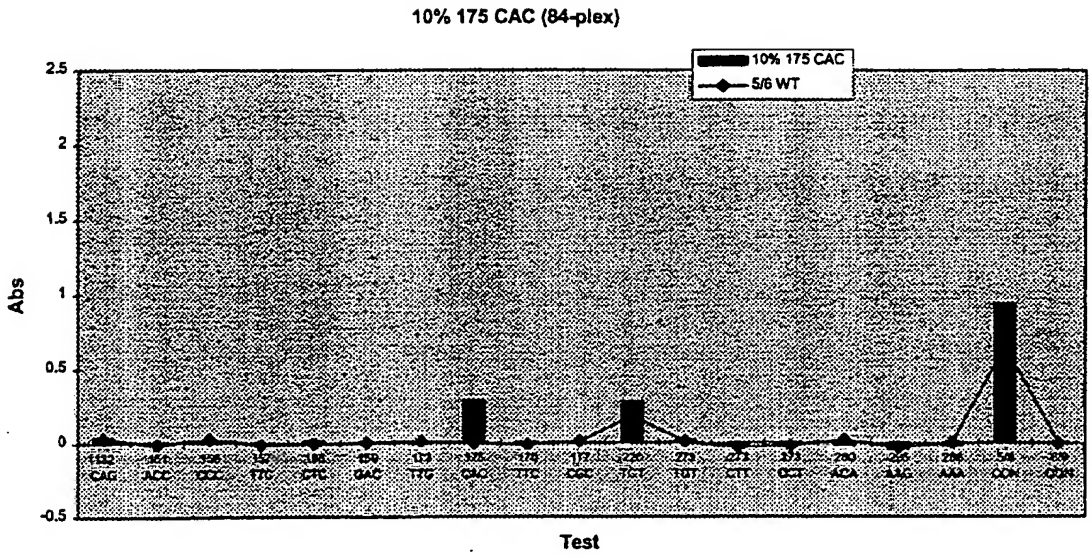


Figure 2b

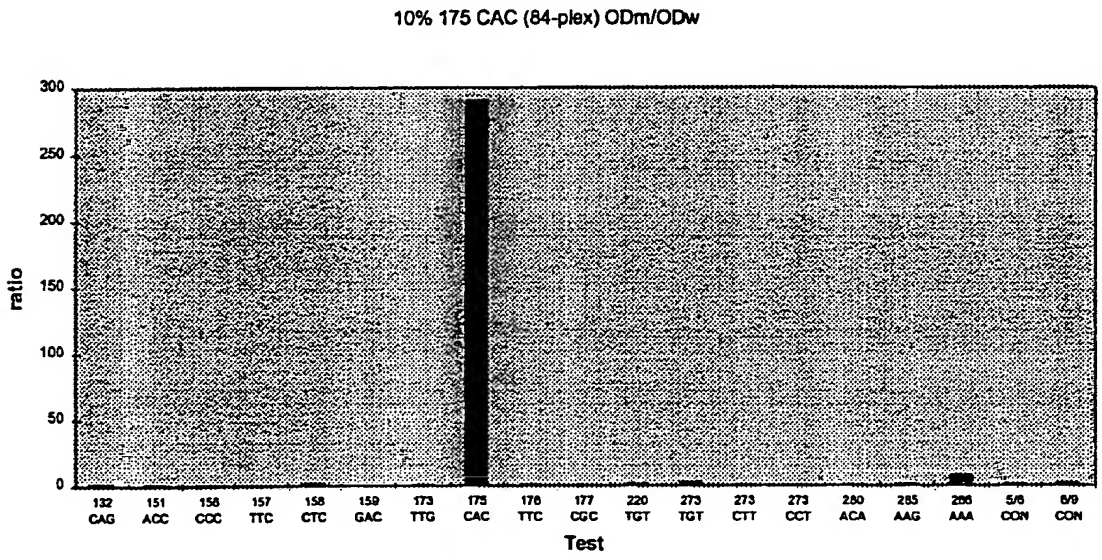


Figure 3a

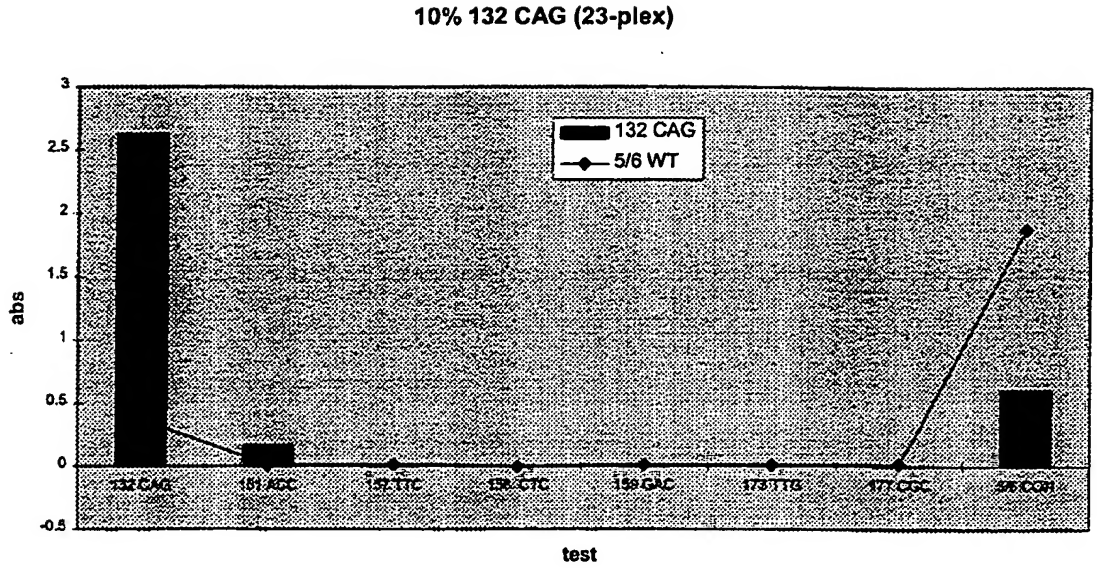


Figure 3b

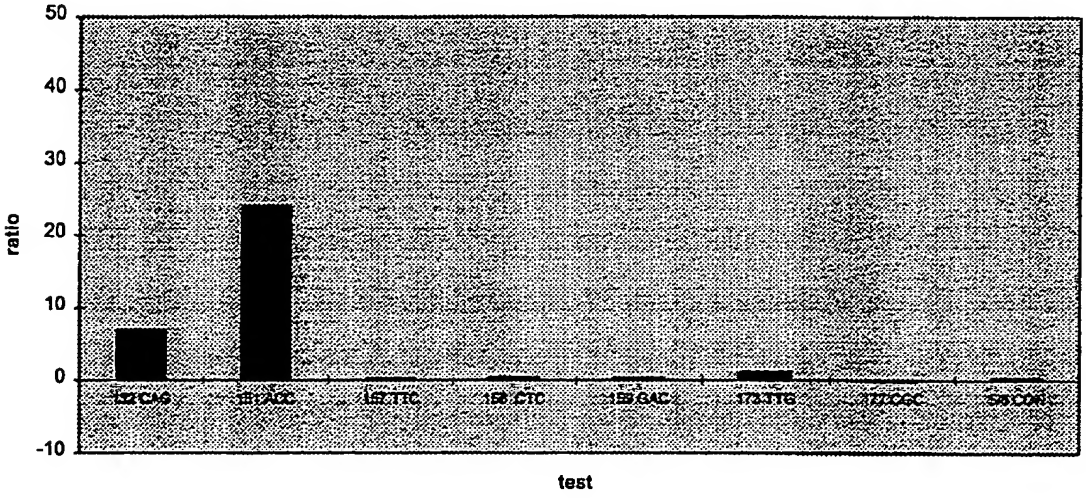


Figure 4a

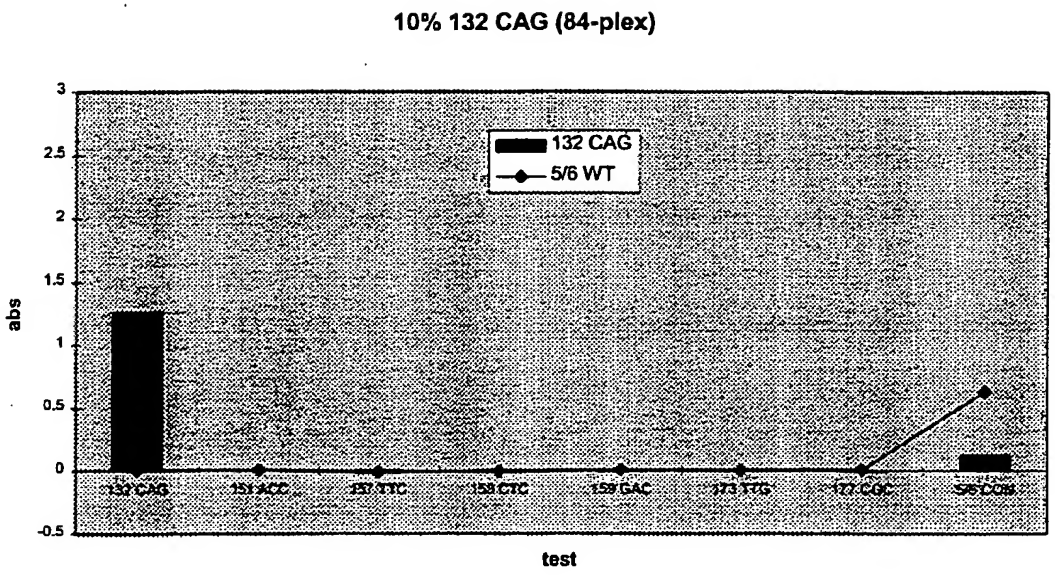
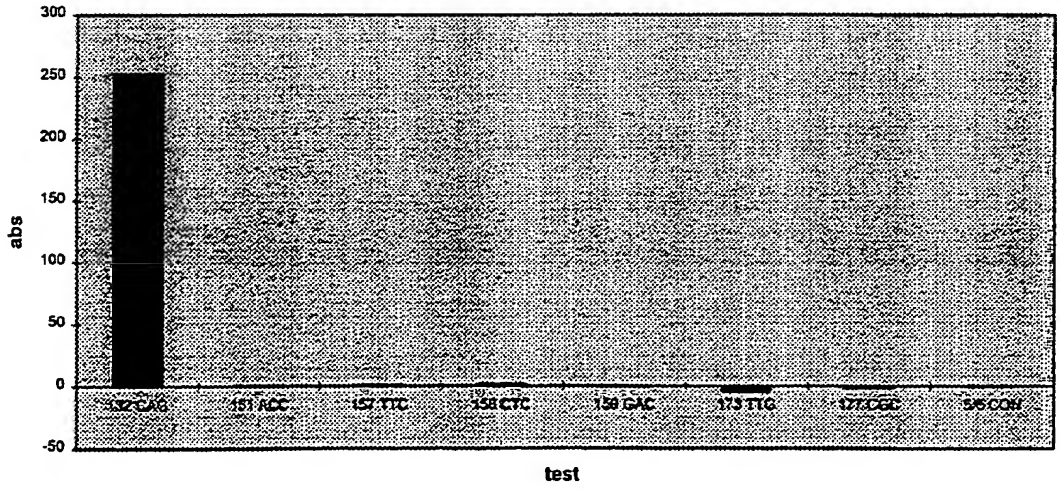


Figure 4b



SEQUENCE LISTING

<110> ZENECA Limited

<120> ASSAY

<130> LDSG/PHM.70477/WO

<140> 9902971.2
<141> 1999-02-11

<150> 9902971.2
<151> 1999-02-11

<160> 3

<170> PatentIn Ver. 2.1

<210> 1
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Single Stranded Capture Sequence.

<400> 1
gctttatgtc cacagatttc 20

<210> 2
<211> 31
<212> DNA
<213> Homo sapiens

<400> 2
atacacagca catgacggag gttgtgagcc a 31

<210> 3
<211> 21

<212> DNA

<213> Homo sapiens

<400> 3

acccggaggg ccactgacaa c 21

INTERNATIONAL SEARCH REPORT

Interna al Application No
PCT/GB 00/00356

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12Q1/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	GB 2 312 747 A (ZENECA LTD) 5 November 1997 (1997-11-05) the whole document	1-9
Y	EP 0 416 817 A (ICI PLC) 13 March 1991 (1991-03-13) cited in the application see whole doc. esp. claims	1-9
X	WO 93 25563 A (HOPE CITY ;WALLACE ROBERT BRUCE (US)) 23 December 1993 (1993-12-23) see whole doc. esp. claims and figures	1-9

-/—

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
30 May 2000	07/06/2000

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Müller, F

1

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Intern al Application No
PCT/GB 00/00356

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
A	OLD J M: "DETECTION OF MUTATIONS BY THE AMPLIFICATION REFRACTORY MUTATION SYSTEM (ARMS)" METHODS IN MOLECULAR BIOLOGY,US,HUMANA PRESS INC., CLIFTON, NJ, vol. 9, 1991, pages 77-84, XP000775640 the whole document	

1

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Information on patent family members

Intern ial Application No
PCT/GB 00/00356

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
GB 2312747 A	05-11-1997	AU 2704997 A	26-11-1997
		EP 0896630 A	17-02-1999
		WO 9742345 A	13-11-1997
EP 0416817 A	13-03-1991	AT 144290 T	15-11-1996
		AU 622349 B	02-04-1992
		AU 6118490 A	14-03-1991
		DE 69028891 D	21-11-1996
		DE 69028891 T	13-03-1997
		GB 2235689 A,B	13-03-1991
		JP 3272686 A	04-12-1991
		NZ 235015 A	26-03-1992
		US 5525494 A	11-06-1996
WO 9325563 A	23-12-1993	CA 2115342 A	23-12-1993
		AU 674127 B	12-12-1996
		AU 2251192 A	04-01-1994
		EP 0607151 A	27-07-1994
		JP 6509946 T	10-11-1994
		US 5981176 A	09-11-1999